

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

CF

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/00, 15/31, 1/20, C07K 1/10, 14/195, A61K 39/00	A1	(11) International Publication Number: WO 97/30149 (43) International Publication Date: 21 August 1997 (21.08.97)
(21) International Application Number: PCT/US97/02547 (22) International Filing Date: 19 February 1997 (19.02.97) (30) Priority Data: 60/011,888 20 February 1996 (20.02.96) US (71) Applicant (for all designated States except US): SMITHKLINE BEECHAM CORPORATION [US/US]; Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): BLACK, Michael, Terrance [GB/US]; 502 Milhouse Way, Chester Springs, PA 19425 (US). BURNHAM, Martin, Karl, Russel [US/US]; 2927 Tanglewood Lane, Norristown, PA 19403 (US). HODGSON, John, Edward [GB/US]; 260 Lapp Road, Malvern, PA 19355 (US). KNOWLES, David, Justin, Charles [GB/GB]; Downsvie House, Redhill, Surrey RH1 6LY (GB). NICHOLAS, Richard, Oakley [GB/GB]; 355 Carmen Drive, Collegeville, PA 19426 (US). PRATT, Julie, M. [GB/GB]; 77 Highfield Drive, Wigston, Leicester LE8 1NP (GB). REICHARD, Raymond, Winfield [US/US]; 2091 Clover Mill Road, Quakertown, PA 18951 (US). ROSENBERG, Martin [US/US]; 241 Mingo Road,	Royersford, PA 19468 (US). WARD, Judith, M. [GB/GB]; Carmela Cottage, 19 Cotmandene, Dorking, Surrey RH4 2BT (GB). (74) Agents: GIMMI, Edward, R. et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, P.O. Box 1539, 709 Swedeland Road, King of Prussia, PA 19406-0939 (US). (81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: NOVEL ERA		
(57) Abstract <p>The invention provides ERA polypeptides and DNA (RNA) encoding ERA polypeptides and methods for producing such polypeptides by recombinant techniques. Also provided are methods for utilizing ERA polypeptide for the protection against infection, particularly bacterial infections.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

Novel ERA

FIELD OF THE INVENTION

This invention relates to newly identified polynucleotides and polypeptides, and their production and uses, as well as their variants, agonists and antagonists, and their uses. In particular, in these and in other regards, the invention relates to novel polynucleotides and polypeptides of the GTP-binding protein family, hereinafter referred to as "ERA".

BACKGROUND OF THE INVENTION

It is particularly preferred to employ Staphylococcal genes and gene products as targets for the development of antibiotics. The Staphylococci make up a medically important genera of microbes. They are known to produce two types of disease, invasive and toxigenic. Invasive infections are characterized generally by abscess formation effecting both skin surfaces and deep tissues. *S. aureus* is the second leading cause of bacteremia in cancer patients. Osteomyelitis, septic arthritis, septic thrombophlebitis and acute bacterial endocarditis are also relatively common. There are at least three clinical conditions resulting from the toxigenic properties of Staphylococci. The manifestation of these diseases result from the actions of exotoxins as opposed to tissue invasion and bacteremia. These conditions include: Staphylococcal food poisoning, scalded skin syndrome and toxic shock syndrome.

A Blast P search result of publicly available sequence data, using the amino acid sequence of the invention as a query sequence, revealed a *Streptococcus mutans* homologue, which is a GTP-binding protein (ERA, sp/P42182/BEX_BACSU BEX PROTEIN).

Clearly, there is a need for factors that may be used to screen compounds for antibiotic activity, such as novel ERA of the invention. These factors may also be used to determine their roles in pathogenesis of infection, dysfunction and disease. There is a further need for identification and characterization of such factors and their antagonists and agonists which can play a role in preventing, ameliorating or correcting infections, dysfunctions or diseases.

The polypeptides of the invention have amino acid sequence homology to a known *Streptococcus mutans* GTP-binding protein ERA protein.

SUMMARY OF THE INVENTION

It is an object of the invention to provide polypeptides that have been identified as novel ERA polypeptides by homology between the amino acid sequence set out in Figure 2

and a known amino acid sequence or sequences of other proteins such as *Streptococcus mutans* GTP-binding protein ERA protein.

It is a further object of the invention to provide polynucleotides that encode ERA polypeptides, particularly polynucleotides that encode the polypeptide herein designated ERA.

5 In a particularly preferred embodiment of this aspect of the invention the polynucleotide comprises a region encoding ERA polypeptides comprising the sequence set out in Figure 1 [SEQ ID NO:1], or a variant thereof.

In another particularly preferred embodiment of the invention there is a novel ERA protein from *Staphylococcus aureus* comprising the amino acid sequence of Figure 2 [SEQ
10 ID NO:2], or a variant thereof.

In accordance with this aspect of the invention there is provided an isolated nucleic acid molecule encoding a mature polypeptide expressible by the *Staphylococcus aureus* WCUH 29

strain contained in NCIMB Deposit No. 40771.

15 In accordance with this aspect of the invention there are provided isolated nucleic acid molecules encoding ERA, particularly *Staphylococcus aureus* ERA, including mRNAs, cDNAs, genomic DNAs. Further embodiments of this aspect of the invention include biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

20 In accordance with another aspect of the invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization. Among the particularly preferred embodiments of this aspect of the invention are naturally occurring allelic variants of ERA and polypeptides encoded thereby.

In accordance with this aspect of the invention there are provided novel polypeptides
25 of *Staphylococcus aureus* referred to herein as ERA as well as biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

Among the particularly preferred embodiments of this aspect of the invention are variants of ERA polypeptide encoded by naturally occurring alleles of the ERA gene.

30 In a preferred embodiment of this aspect of the invention there are provided methods for producing the aforementioned ERA polypeptides.

In accordance with yet another aspect of the invention, there are provided inhibitors to such polypeptides, useful as antibacterial agents, including, for example, antibodies.

In accordance with certain preferred embodiments of this aspect of the invention, there are provided products, compositions and methods for (i) assessing ERA expression, (ii) treating disease, for example, disease, such as, infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic absces, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, dacryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and perinephric absces, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis), (iii) assaying genetic variation, (iv) and administering a ERA polypeptide or polynucleotide to an organism to raise an immunological response against a bacteria, especially a *Staphylococcus aureus* bacteria.

In accordance with certain preferred embodiments of this and other aspects of the invention there are provided polynucleotides that hybridize to ERA polynucleotide sequences, particularly under stringent conditions.

In certain preferred embodiments of this aspect of the invention there are provided antibodies against ERA polypeptides.

In accordance with another aspect of the invention, there are provided ERA agonists and antagonists each of which are also preferably bacteriostatic or bacteriocidal.

In a further aspect of the invention there are provided compositions comprising a ERA polynucleotide or a ERA polypeptide for administration to a cell or to a multicellular organism.

Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings depict certain embodiments of the invention. They are illustrative only and do not limit the invention otherwise disclosed herein.

Figure 1 shows the polynucleotide sequence of *Staphylococcus aureus* ERA [SEQ ID NO:1].

Figure 2 shows the amino acid sequence of *Staphylococcus aureus* ERA [SEQ ID NO:2] deduced from the polynucleotide sequence of Figure 1.

GLOSSARY

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"Host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous polynucleotide sequence.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity and similarity between two sequences, both terms are well known to skilled artisans (*Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Methods commonly employed to determine identity or similarity between sequences include, but are not limited to those disclosed in Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48:1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990)). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990)).

"Isolated" means altered "by the hand of man" from its natural state, i.e., if it occurs in nature, it has been changed or removed from its original environment, or both. For

example, polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide(s)" generally refers to any polyribonucleotide or
5 polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotide(s)" include, without limitation, single-and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or,
10 more typically, double-stranded, or triple-stranded, or a mixture of single- and double-stranded regions. In addition, polynucleotide as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of
15 the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term "polynucleotide(s)" includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotide(s)" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name
20 just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term "polynucleotide(s)" as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and
25 cells, including, for example, simple and complex cells. "Polynucleotide(s)" embraces short polynucleotides often referred to as oligonucleotide(s).

"Polypeptide(s)" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds. "Polypeptide(s)" refers to both short chains, commonly referred to as peptides, oligopeptides and oligomers and to
30 longer chains generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene encoded amino acids. "Polypeptide(s)" include those modified either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques which are well known to the art. Such modifications are

well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. See, for instance, *PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993) and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in *POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS*, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., *Meth. Enzymol.* 182:626-646 (1990) and Rattan et al., *Protein Synthesis: Posttranslational Modifications and Aging*, Ann. N.Y. Acad. Sci. 663: 48-62 (1992). Polypeptides may be branched or cyclic, with or without branching. Cyclic, branched and branched circular polypeptides may result from post-translational natural processes and may be made by entirely synthetic methods, as well.

"Variant(s)" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from

another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques, by direct synthesis, and by other recombinant methods known to skilled artisans.

10 DESCRIPTION OF THE INVENTION

The invention relates to novel ERA polypeptides and polynucleotides as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of a novel ERA gene of *Staphylococcus aureus*, which is related by amino acid sequence homology to *Streptococcus mutans* GTP-binding protein ERA polypeptide. The invention relates especially to ERA having the nucleotide and amino acid sequences set out in Figure 1 and Figure 2 respectively, and to the ERA nucleotide sequences of the DNA in NCIMB Deposit No. 40771 and amino acid sequences encoded thereby.

Techniques are available to evaluate temporal gene expression in bacteria, particularly as it applies to viability under laboratory and host infection conditions. A number of methods can be used to identify genes which are essential to survival *per se*, or essential to the establishment and/or maintenance of an infection. Identification of expression of a sequence by one of these methods yields additional information about its function and assists in the selection of such sequence for further development as a screening target. Briefly, these approaches include for example;

25 1) Signature Tagged Mutagenesis (STM)

This technique is described by Hensel *et al.*, *Science* 269: 400-403(1995), the contents of which is incorporated by reference for background purposes. Signature tagged mutagenesis identifies genes necessary for the establishment/maintenance of infection in a given infection model.

30 The basis of the technique is the random mutagenesis of target organism by various means (e.g., transposons) such that unique DNA sequence tags are inserted in close proximity to the site of mutation. The tags from a mixed population of bacterial mutants and bacteria recovered from an infected hosts are detected by amplification, radiolabeling

and hybridization analysis. Mutants attenuated in virulence are revealed by absence of the tag from the pool of bacteria recovered from infected hosts.

In *Staphylococcus aureus*, because the transposon system is less well developed, a more efficient way of creating the tagged mutants is to use the insertion-duplication
5 mutagenesis technique as described by Morrison *et al.*, *J. Bacteriol.* 159:870 (1984) the contents of which is incorporated by reference for background purposes.

2) In Vivo Expression Technology (IVET)

This technique is described by Camilli *et al.*, *Proc. Nat'l. Acad. Sci. USA.* 91:2634-2638 (1994) and Mahan *et al.*, *Infectious Agents and Diseases* 2:263-268 (1994), the
10 contents of each of which is incorporated by reference for background purposes. IVET identifies genes up-regulated during infection when compared to laboratory cultivation, implying an important role in infection. Sequences identified by this technique are implied to have a significant role in infection establishment/maintenance.

In this technique random chromosomal fragments of target organism are cloned
15 upstream of a promoter-less reporter gene in a plasmid vector. The pool is introduced into a host and at various times after infection bacteria may be recovered and assessed for the presence of reporter gene expression. The chromosomal fragment carried upstream of an expressed reporter gene should carry a promoter or portion of a gene normally upregulated during infection. Sequencing upstream of the reporter gene allows identification of the up
20 regulated gene.

3) Differential display

This technique is described by Chuang *et al.*, *J. Bacteriol.* 175:2026-2036 (1993), the contents of which is incorporated by reference for background purposes. This method identifies those genes which are expressed in an organism by identifying mRNA present
25 using randomly-primed RT-PCR. By comparing pre-infection and post infection profiles, genes up and down regulated during infection can be identified and the RT-PCR product sequenced and matched to library sequences.

4) Generation of conditional lethal mutants by transposon mutagenesis.

This technique, described by de Lorenzo, V. *et al.*, *Gene* 123:17-24 (1993); Neuwald, A. F. *et al.*, *Gene* 125: 69-73(1993); and Takiff, H. E. *et al.*, *J. Bacteriol.* 174:1544-1553(1992), the contents of which is incorporated by reference for

5 background purposes, identifies genes whose expression are essential for cell viability.

In this technique transposons carrying controllable promoters, which provide transcription outward from the transposon in one or both directions, are generated. Random insertion of these transposons into target organisms and subsequent isolation of insertion mutants in the presence of inducer of promoter activity ensures that insertions
10 which separate promoter from coding region of a gene whose expression is essential for cell viability will be recovered. Subsequent replica plating in the absence of inducer identifies such insertions, since they fail to survive. Sequencing of the flanking regions of the transposon allows identification of site of insertion and identification of the gene disrupted. Close monitoring of the changes in cellular processes/morphology during growth in the
15 absence of inducer yields information on likely function of the gene. Such monitoring could include flow cytometry (cell division, lysis, redox potential, DNA replication), incorporation of radiochemically labeled precursors into DNA, RNA, protein, lipid, peptidoglycan, monitoring reporter enzyme gene fusions which respond to known cellular stresses.

20 5) Generation of conditional lethal mutants by chemical mutagenesis.

This technique is described by Beckwith, J., *Methods in Enzymology* 204: 3-18(1991), the contents of which are incorporated herein by reference for background purposes. In this technique random chemical mutagenesis of target organism, growth at temperature other than physiological temperature (permissive temperature) and subsequent
25 replica plating and growth at different temperature (e.g., 42°C to identify ts, 25°C to identify cs) are used to identify those isolates which now fail to grow (conditional mutants). As above close monitoring of the changes upon growth at the non-permissive temperature yields information on the function of the mutated gene. Complementation of conditional lethal mutation by library from target organism and sequencing of complementing gene
30 allows matching with library sequences.

Each of these techniques may have advantages or disadvantage depending on the particular application. The skilled artisan would choose the approach that is the most relevant with the particular end use in mind. For example, some genes might be recognised

as essential for infection but in reality are only necessary for the initiation of infection and so their products would represent relatively unattractive targets for antibacterials developed to cure established and chronic infections.

6) RT-PCR

5 Bacterial messenger RNA, preferably that of *Staphylococcus aureus*, is isolated from bacterial infected tissue, e.g., 48 hour murine lung infections, and the amount of each mRNA species assessed by reverse transcription of the RNA sample primed with random hexanucleotides followed by PCR with gene specific primer pairs. The determination of the presence and amount of a particular mRNA species by quantification of the resultant
10 PCR product provides information on the bacterial genes which are transcribed in the infected tissue. Analysis of gene transcription can be carried out at different times of infection to gain a detailed knowledge of gene regulation in bacterial pathogenesis allowing for a clearer understanding of which gene products represent targets for screens for novel antibacterials. Because of the gene specific nature of the PCR primers employed it should
15 be understood that the bacterial mRNA preparation need not be free of mammalian RNA. This allows the investigator to carry out a simple and quick RNA preparation from infected tissue to obtain bacterial mRNA species which are very short lived in the bacterium (in the order of 2 minute half-lives). Optimally the bacterial mRNA is prepared from infected murine lung tissue by mechanical disruption in the presence of TRIzol
20 (GIBCO-BRL) for very short periods of time, subsequent processing according to the manufacturers of TRIzol reagent and DNAase treatment to remove contaminating DNA. Preferably the process is optimized by finding those conditions which give a maximum amount of bacterial 16S ribosomal RNA, preferably that of *Staphylococcus aureus*, as detected by probing Northern blots with a suitably labeled sequence specific oligonucleotide
25 probe. Typically, a 5' dye labelled primer is used in each PCR primer pair in a PCR reaction which is terminated optimally between 8 and 25 cycles. The PCR products are separated on 6% polyacrylamide gels with detection and quantification using GeneScanner (manufactured by ABI).

Use of the of these technologies when applied to the sequences of the invention
30 enables identification of bacterial proteins expressed during infection, inhibitors of which would have utility in anti-bacterial therapy.

Deposited materials

A deposit containing a *Staphylococcus aureus* WCUH 29 strain has been deposited with the National Collections of Industrial and Marine Bacteria Ltd. (NCIMB), 23 St. Machar Drive, Aberdeen AB2 1RY, Scotland on 11 September 1995 and assigned NCIMB Deposit No. 40771. The *Staphylococcus aureus* strain deposit is referred to herein as "the deposited strain" or as "the DNA of the deposited strain."

The deposited material is a strain that contains the full length ERA DNA, referred to as "NCIMB 40771" upon deposit.

The sequence of the polynucleotides contained in the deposited material, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

The deposit has been made under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposit is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that required under 35 U.S.C. §112.

A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

Polypeptides

The polypeptides of the invention include the polypeptide of Figure 2 [SEQ ID NO:2] (in particular the mature polypeptide) as well as polypeptides and fragments, such as the polypeptide of SEQ ID NO:4, and particularly those polypeptides which have the biological activity of ERA, and also those which have at least 70% identity to the polypeptide of Figure 2 [SEQ ID NO:2] or Figure 4 [SEQ ID NO:4] or the relevant portion, preferably at least 80% identity to the polypeptide of Figure 2 [SEQ ID NO:2] or Figure 4 [SEQ ID NO:4], and more preferably at least 90% similarity (more preferably at least 90% identity) to the polypeptide of Figure 2 [SEQ ID NO:2] or Figure 4 [SEQ ID NO:4] and still more preferably at least 95% similarity (still more preferably at least 95% identity) to the polypeptide of Figure 2 [SEQ ID NO:2] or Figure 4 [SEQ ID NO:4] and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 amino acids and more preferably at least 50 amino acids.

Variants that are fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these

variants may be employed as intermediates for producing the full-length polypeptides. Variants that are fragments of the polynucleotides of the invention may be used to synthesize full-length polynucleotides of the invention.

5 A fragment is a variant polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned polypeptides. As with ERA polypeptides fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region, a single larger polypeptide.

10 Preferred fragments include, for example, truncation polypeptides having a portion of the amino acid sequence of Figure 2 [SEQ ID NO:2], or of variants thereof, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Degradation forms of the polypeptides of the invention in a host cell, particularly a
15 *Staphylococcus aureus*, are also preferred. Also preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding
20 region, and high antigenic index regions.

Also preferred are biologically active fragments which are those fragments that mediate activities of ERA, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those fragments that are antigenic or immunogenic in an animal, especially in a human.

25 **Polynucleotides**

Another aspect of the invention relates to isolated polynucleotides which encode the ERA polypeptide having the deduced amino acid sequence of Figure 2 [SEQ ID NO:2] and polynucleotides closely related thereto and variants thereto.

Using the information provided herein, such as the polynucleotide sequence set out in
30 Figure 1 [SEQ ID NO:1], a polynucleotide of the invention encoding ERA polypeptide may be obtained using standard cloning and screening, such as those for cloning and sequencing chromosomal DNA fragments from *Staphylococcus aureus* WCUH 29 cells as starting material, followed by obtaining a full length clone. For example, to obtain a polynucleotide

sequence of the invention, such as that sequence given in Figure 1 [SEQ ID NO:1] or Figure 3 [SEQ ID NO:3], typically a library of clones of chromosomal DNA of *Staphylococcus aureus* WCUH 29 in *E.coli* or some other suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent conditions. By sequencing the individual clones thus identified with sequencing primers designed from the original sequence it is then possible to extend the sequence in both directions to determine the full gene sequence. Conveniently such sequencing is performed using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). (see Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). Illustrative of the invention, the polynucleotide set out in Figure 1 [SEQ ID NO:1] was discovered in a DNA library derived from *Staphylococcus aureus* WCUH 29.

The DNA sequence thus obtained is set out in Figure 1 [SEQ ID NO:1]. It contains an open reading frame encoding a protein having about the number of amino acid residues set forth in Figure 2 [SEQ ID NO:2] with a deduced molecular weight that can be calculated using amino acid residue molecular weight values well known in the art.

ERA of the invention is structurally related to other proteins of the GTP-binding protein family, as shown by the results of sequencing the DNA encoding ERA of the deposited strain. The protein exhibits greatest homology to Streptococcus mutans GTP-binding protein ERA protein among known proteins. ERA of Figure 2 [SEQ ID NO:2] has about 61% identity over its entire length and about 79% similarity over its entire length with the amino acid sequence of Streptococcus mutans GTP-binding protein ERA.

Sequence of the invention may also be identical over its entire length to the coding sequence in Figure 1 [SEQ ID NO:1] or Figure 3 [SEQ ID NO:3]. The sequence of Figure 3 [SEQ ID NO:3], from nucleotides 242 to 523, putatively encodes the polypeptide of Figure 4 [SEQ ID NO:4].

Also provided by the invention is the coding sequence for the mature polypeptide or a fragment thereof, by itself as well as the coding sequence for the mature polypeptide or a fragment in reading frame with other coding sequence, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence. The polynucleotide may also

contain non-coding sequences, including for example, but not limited to non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences, termination signals, ribosome binding sites, sequences that stabilize mRNA, introns, polyadenylation signals, and additional coding sequence which encode additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc. Natl. Acad. Sci., USA* 86: 821-824 (1989), or an HA tag (Wilson *et al.*, *Cell* 37: 767 (1984)). Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

In accordance with the foregoing, the term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides which include a sequence encoding a polypeptide of the invention, particularly bacterial, and more particularly the *Staphylococcus aureus* ERA having the amino acid sequence set out in Figure 2 [SEQ ID NO:2]. The term encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, interrupted by integrated phage or an insertion sequence or editing) together with additional regions, that also may contain coding and/or non-coding sequences.

The invention further relates to variants of the herein above described polynucleotides which encode for variants of the polypeptide having the deduced amino acid sequence of Figure 2 [SEQ ID NO:2].

Further particularly preferred embodiments are polynucleotides encoding ERA variants, which have the amino acid sequence of ERA polypeptide of Figure 2 [SEQ ID NO:2] in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of ERA.

Further preferred embodiments of the invention are polynucleotides that are at least 70% identical over their entire length to a polynucleotide encoding ERA polypeptide having the amino acid sequence set out in Figure 2 [SEQ ID NO:2], and polynucleotides which are complementary to such polynucleotides. Alternatively, most highly preferred are polynucleotides that comprise a region that is at least 80% identical over their entire length to a polynucleotide encoding ERA polypeptide of the *Staphylococcus aureus* DNA of the deposited strain and polynucleotides complementary thereto. In this regard, polynucleotides at least 90% identical over their entire length to the same are particularly preferred, and

among these particularly preferred polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

5 Preferred embodiments in this respect, moreover, are polynucleotides which encode polypeptides which retain substantially the same biological function or activity as the mature polypeptide encoded by the DNA of Figure 1 [SEQ ID NO:1].

The invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the invention especially relates to polynucleotides which
10 hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. An example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl,
15 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C. Hybridization and wash conditions are well known and exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein, the
20 disclosure of which is hereby incorporated in its entirety by reference.

The invention also provides a polynucleotide consisting essentially of a polynucleotide sequence obtainable by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in SEQ ID NO:1 under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence
25 set forth in SEQ ID NO:1 or a fragment thereof; and isolating said DNA sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers described elsewhere herein.

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, polynucleotides of the invention as discussed above, may be used as a hybridization
30 probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding ERA and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to the ERA gene. Such probes generally will comprise at least 15 bases.

Preferably, such probes will have at least 30 bases and may have at least 50 bases. Particularly preferred probes will have at least 30 bases and will have 50 bases or less.

For example, the coding region of the ERA gene may be isolated by screening using the known DNA sequence to synthesize an oligonucleotide probe. A labeled oligonucleotide
5 having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The polynucleotides and polypeptides of the invention may be employed as research reagents and materials for discovery of treatments of and diagnostics for disease, particularly
10 human disease, as further discussed herein relating to polynucleotide assays, *inter alia*.

Polynucleotides of the invention that are oligonucleotides derived from the sequences of SEQ ID NOS:1 and 2 may be used in the processes herein as described, but preferably for PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in infected tissue. It is recognized that such sequences will
15 also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The polynucleotides may encode a polypeptide which is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such
20 sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

25 A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In sum, a polynucleotide of the invention may encode a mature protein, a mature
30 protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence

and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

Vectors, host cells, expression

The invention also relates to vectors which comprise a polynucleotide or
5 polynucleotides of the invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

For recombinant production, host cells can be genetically engineered to incorporate
10 expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis et al., *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as. calcium
15 phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, *E. coli*, streptomyces and *Bacillus subtilis* cells; fungal cells, such
20 as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used to produce a polypeptide of the invention. Such vectors include, among others, chromosomal, episomal and virus-derived
25 vectors, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic
30 elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be

inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, (*supra*).

For secretion of the translated protein into the lumen of the endoplasmic reticulum,
5 into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid
10 extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during
15 isolation and or purification.

Diagnostic Assays

This invention is also related to the use of the ERA polynucleotides of the invention for use as diagnostic reagents. Detection of ERA in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of a disease. Eukaryotes
20 (herein also "individual(s)"), particularly mammals, and especially humans, infected with an organism comprising the ERA gene may be detected at the DNA level by a variety of techniques.

Nucleic acids for diagnosis may be obtained from an infected individual's cells and tissues, such as bone, blood, muscle, cartilage, and skin. Genomic DNA may be used directly
25 for detection or may be amplified enzymatically by using PCR or other amplification technique prior to analysis. RNA or cDNA may also be used in the same ways. Using amplification, characterization of the strain of prokaryote present in a eukaryote, particularly a mammal, and especially a human, may be made by an analysis of the genotype of the prokaryote gene. Deletions and insertions can be detected by a change in size of the amplified
30 product in comparison to the genotype of a reference sequence. Point mutations can be identified by hybridizing amplified DNA to labeled ERA polynucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by

alterations in the electrophoretic mobility of the DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, e.g., Myers et al., *Science*, 230: 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or a chemical cleavage method. See, e.g., Cotton et al., *Proc. Natl. Acad. Sci., USA*, 85: 4397-4401 (1985).

Cells carrying mutations or polymorphisms in the gene of the invention may also be detected at the DNA level by a variety of techniques, to allow for serotyping, for example. For example, RT-PCR can be used to detect mutations. It is particularly preferred to use RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA or cDNA may also be used for the same purpose, PCR or RT-PCR. As an example, PCR primers complementary to the nucleic acid encoding ERA can be used to identify and analyze mutations. These primers may also be used for amplifying ERA DNA isolated from a sample derived from an individual. The invention further provides these primers with 1, 2, 3 or 4 nucleotides removed from the 5' and/or the 3' end. The primers may be used to amplify the gene isolated from an infected individual such that the gene may then be subject to various techniques for elucidation of the DNA sequence. In this way, mutations in the DNA sequence may be detected and used to diagnose infection and to serotype or classify the infectious agent.

The invention provides a process for diagnosing, disease, preferably bacterial infections, more preferably infections by *Staphylococcus aureus*, and most preferably disease, such as, infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic abscess, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, dacryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and perinephric abscess, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis), comprising determining from a sample derived from an individual a increased level of expression of polynucleotide having the sequence of Figure 1 [SEQ ID NO: 1]. Increased or decreased expression of ERA polynucleotide can be measured using any one of the methods well known in the art for the quantitation of polynucleotides, such as, for example, amplification, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods.

In addition, a diagnostic assay in accordance with the invention for detecting over-expression of ERA protein compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used to determine levels of a ERA protein, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Antibodies

The polypeptides of the invention or variants thereof, or cells expressing them can be used as an immunogen to produce antibodies immunospecific for such polypeptides. "Antibodies" as used herein includes monoclonal and polyclonal antibodies, chimeric, single chain, simianized antibodies and humanized antibodies, as well as Fab fragments, including the products of an Fab immunoglobulin expression library.

Antibodies generated against the polypeptides of the invention can be obtained by administering the polypeptides or epitope-bearing fragments, analogues or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique known in the art which provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pg. 77-96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies.

Alternatively phage display technology could be utilized to select antibody genes with binding activities towards the polypeptide either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-Fbp or from naive libraries (McCafferty, J. *et al.*, (1990), *Nature* 348, 552-554; Marks, J. *et al.*, (1992) *Biotechnology* 10, 779-783). The affinity of these antibodies can also be improved by chain shuffling (Clackson, T. *et al.*, (1991) *Nature* 352, 624-628).

If two antigen binding domains are present each domain may be directed against a different epitope - termed 'bispecific' antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides to purify the polypeptides by affinity chromatography.

Thus, among others, antibodies against ERA may be employed to treat infections, particularly bacterial infections and especially disease, such as, infections of the upper
5 respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower
respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis),
gastrointestinal (e.g., secretory diarrhoea, splenic absces, retroperitoneal abscess), CNS (e.g.,
cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and
orbital cellulitis, dacryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and
10 perinephric absces, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous
abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis,
osteomyelitis).

Polypeptide variants include antigenically, epitopically or immunologically
equivalent variants which form a particular aspect of this invention. The term
15 "antigenically equivalent derivative" as used herein encompasses a polypeptide or its
equivalent which will be specifically recognised by certain antibodies which, when raised
to the protein or polypeptide according to the invention, interfere with the immediate
physical interaction between pathogen and mammalian host. The term "immunologically
equivalent derivative" as used herein encompasses a peptide or its equivalent which when
20 used in a suitable formulation to raise antibodies in a vertebrate, the antibodies act to
interfere with the immediate physical interaction between pathogen and mammalian host.

The polypeptide, such as an antigenically or immunologically equivalent derivative
or a fusion protein thereof is used as an antigen to immunize a mouse or other animal such
as a rat or chicken. The fusion protein may provide stability to the polypeptide. The
25 antigen may be associated, for example by conjugation, with an immunogenic carrier
protein for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH).
Alternatively a multiple antigenic peptide comprising multiple copies of the protein or
polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be
sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

30 Preferably the antibody or variant thereof is modified to make it less immunogenic
in the individual. For example, if the individual is human the antibody may most
preferably be "humanized"; where the complementarity determining region(s) of the
hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for

example as described in Jones, P. et al. (1986), Nature 321, 522-525 or Tempest et al., (1991) Biotechnology 9, 266-273.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff et al., Hum Mol Genet 1992, 1:363, Manthorpe et al., Hum. Gene Ther. 1963:4, 419), delivery of DNA complexed with specific protein carriers (Wu et al., J Biol Chem 1989:264,16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, PNAS, 1986:83,9551), encapsulation of DNA in various forms of liposomes (Kaneda et al., Science 1989:243,375), particle bombardment (Tang et al., Nature 1992, 356:152, Eisenbraun et al., DNA Cell Biol 1993, 12:791) and *in vivo* infection using cloned retroviral vectors (Seeger et al., PNAS 1984:81,5849).

Polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, *e.g.*, Coligan *et al.*, *Current Protocols in Immunology 1(2)*: Chapter 5 (1991).

Antagonists and agonists - assays and molecules

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of ERA polypeptides or polynucleotides.

For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, comprising ERA polypeptide and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule which may be a ERA agonist or antagonist. The ability of the candidate molecule to agonize or antagonize the ERA polypeptide is reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Molecules which bind gratuitously, *i.e.*, without inducing the effects of ERA are most likely to be good antagonists. Molecules that bind well and increase the rate of product production from substrate are agonists. The rate or level of production of product from substrate may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric labeled substrate converted into product, a reporter gene that is responsive to changes in ERA activity, and binding assays known in the art.

Another example of an assay for ERA antagonists is a competitive assay that combines ERA and a potential antagonist with ERA-binding molecules, recombinant ERA binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. ERA can be labeled, such as by
5 radioactivity or a colorimetric compound, such that the number of ERA molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polypeptide of the invention and thereby inhibit or extinguish its
10 activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing ERA-induced activities, thereby preventing the action of ERA by excluding ERA from binding.

Potential antagonists include a small molecule which binds to and occupies the
15 binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, *J. Neurochem.* 56: 560 (1991); *OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION*,
20 CRC Press, Boca Raton, FL (1988), for a description of these molecules). Preferred potential antagonists include compounds related to and variants of ERA.

In a particular aspect the invention provides the use of the polypeptide, polynucleotide or inhibitor of the invention to interfere with the initial physical interaction between a pathogen and mammalian host responsible for sequelae of infection. In
25 particular the molecules of the invention may be used: i) in the prevention of adhesion of bacteria, in particular gram positive bacteria, to mammalian extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; ii) to block ERA protein mediated mammalian cell invasion by, for example, initiating phosphorylation of mammalian tyrosine kinases (Rosenshine *et al.*, *Infect. Immun.* 60:2211 (1992); iii) to block
30 bacterial adhesion between mammalian extracellular matrix proteins and bacterial ERA proteins which mediate tissue damage; iv) to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques.

Each of the DNA sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein upon expression can be used as a target for the screening of antibacterial drugs. Additionally, the DNA sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The antagonists and agonists may be employed for instance to inhibit disease, such as, infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic abscess, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, dacryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and perinephric abscess, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis).

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal which comprises inoculating the individual with ERA, or a fragment or variant thereof, adequate to produce antibody to protect said individual from infection, particularly bacterial infection and most particularly *Staphylococcus aureus* infections. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises, through gene therapy, delivering gene encoding ERA, or a fragment or a variant thereof, for expressing ERA, or a fragment or a variant thereof *in vivo* in order to induce an immunological response to produce antibody to protect said individual from disease.

A further aspect of the invention relates to an immunological composition which, when introduced into a host capable or having induced within it an immunological response, induces an immunological response in such host to a ERA or protein coded therefrom, wherein the composition comprises a recombinant ERA or protein coded therefrom comprising DNA which codes for and expresses an antigen of said ERA or protein coded therefrom.

The ERA or a fragment thereof may be fused with co-protein which may not by itself produce antibodies, but is capable of stabilizing the first protein and producing a

fused protein which will have immunogenic and protective properties. Thus fused recombinant protein, preferably further comprises an antigenic co-protein, such as Glutathione-S-transferase (GST) or beta-galactosidase, relatively large co-proteins which solubilise the protein and facilitate production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system. The co-protein may be attached to either the amino or carboxy terminus of the first protein.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. *et al.* Science 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof which have been shown to encode non-variable regions of bacterial cell surface proteins in DNA constructs used in such genetic immunization experiments in animal models of infection with *Staphylococcus aureus* will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value from the requisite organ of the animal successfully resisting or clearing infection for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *Staphylococcus aureus* infections, in mammals, particularly humans.

The polypeptide may be used as an antigen for vaccination of a host to produce specific antibodies which protect against invasion of bacteria, for example by blocking adherence of bacteria to damaged tissue. Examples of tissue damage include wounds in skin or connective tissue caused e.g. by mechanical, chemical or thermal damage or by implantation of indwelling devices, or wounds in the mucous membranes, such as the mouth, mammary glands, urethra or vagina.

The invention also includes a vaccine formulation which comprises the immunogenic recombinant protein together with a suitable carrier. Since the protein may be broken down in the stomach, it is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and

solutes which render the formulation isotonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

While the invention has been described with reference to certain ERA, it is to be understood that this covers fragments of the naturally occurring protein and similar proteins with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant protein.

Compositions, kits and administration

The invention also relates to compositions comprising the polynucleotide or the polypeptides discussed above or the agonists or antagonists. The polypeptides of the invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides and other compounds of the invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively the composition may be formulated for topical application for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug
5 penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such carriers may constitute from about 1% to about 98% by weight of the formulation; more usually they will constitute up to about 80% by weight of the formulation.

10 For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of
15 course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

In-dwelling devices include surgical implants, prosthetic devices and catheters, i.e., devices that are introduced to the body of an individual and remain in position for an extended time. Such devices include, for example, artificial joints, heart valves,
20 pacemakers, vascular grafts, vascular catheters, cerebrospinal fluid shunts, urinary catheters, continuous ambulatory peritoneal dialysis (CAPD) catheters, etc.

The composition of the invention may be administered by injection to achieve a systemic effect against relevant bacteria shortly before insertion of an in-dwelling device. Treatment may be continued after surgery during the in-body time of the device. In
25 addition, the composition could also be used to broaden perioperative cover for any surgical technique to prevent bacterial wound infections, especially *Staphylococcus aureus* wound infections.

Many orthopaedic surgeons consider that humans with prosthetic joints should be considered for antibiotic prophylaxis before dental treatment that could produce a
30 bacteremia. Late deep infection is a serious complication sometimes leading to loss of the prosthetic joint and is accompanied by significant morbidity and mortality. It may therefore be possible to extend the use of the active agent as a replacement for prophylactic antibiotics in this situation.

In addition to the therapy described above, the compositions of this invention may be used generally as a wound treatment agent to prevent adhesion of bacteria to matrix proteins exposed in wound tissue and for prophylactic use in dental treatment as an alternative to, or in conjunction with, antibiotic prophylaxis.

5 Alternatively, the composition of the invention may be used to bathe an indwelling device immediately before insertion. The active agent will preferably be present at a concentration of 1µg/ml to 10mg/ml for bathing of wounds or indwelling devices.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is
10 0.5-5µg/kg of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks. With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention which would preclude their administration to suitable individuals.

The antibodies described above may also be used as diagnostic reagents to detect
15 the presence of bacteria containing the ERA protein.

EXAMPLES

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples are illustrative, but do not limit the invention.

20 Example 1 Library Production

The polynucleotide having the DNA sequence given in SEQ ID NO:1 was obtained from a library of clones of chromosomal DNA of *Staphylococcus aureus* in *E. coli*. In some cases the sequencing data from two or more clones containing overlapping
25 *Staphylococcus aureus* DNAs was used to construct the contiguous DNA sequence in SEQ ID NO:1. Libraries may be prepared by routine methods, for example:
Methods 1 and 2 below.

Total cellular DNA is isolated from *Staphylococcus aureus* WCUH 29 according to standard procedures and size-fractionated by either of two methods.

Method 1

30 Total cellular DNA is mechanically sheared by passage through a needle in order to size-fractionate according to standard procedures. DNA fragments of up to 11kbp in size are rendered blunt by treatment with exonuclease and DNA polymerase, and EcoRI linkers added. Fragments are ligated into the vector Lambda ZapII that has been cut with EcoRI,

the library packaged by standard procedures and *E.coli* infected with the packaged library. The library is amplified by standard procedures.

Method 2

Total cellular DNA is partially hydrolyzed with a one or a combination of
5 restriction enzymes appropriate to generate a series of fragments for cloning into library
vectors (e.g., *RsaI*, *Pall*, *AluI*, *BshI235I*), and such fragments are size-fractionated
according to standard procedures. *EcoRI* linkers are ligated to the DNA and the fragments
then ligated into the vector Lambda ZapII that have been cut with *EcoRI*, the library
packaged by standard procedures, and *E.coli* infected with the packaged library. The library
10 is amplified by standard procedures.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Black, Michael

Burnham, Martin
Hodgson, John
Knowles, David
Nicholas, Richard
Pratt, Julie
Reichard, Raymond
Rosenberg, Martin
Ward, Judith

(ii) TITLE OF THE INVENTION: Novel ERA

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SmithKline Beecham Corporation
(B) STREET: 709 Swedeland Road
(C) CITY: King of Prussia
(D) STATE: PA
(E) COUNTRY:
(F) ZIP: 19406

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: 19-FEB-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/011,888
(B) FILING DATE: 20-FEB-1996

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Gimmi, Edward R
- (B) REGISTRATION NUMBER: 38,891
- (C) REFERENCE/DOCKET NUMBER: GM50006-1

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 610-270-4478
- (B) TELEFAX: 610-270-4026
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 900 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

ATGACAGAAC ATAAATCAGG ATTTGTTTCA ATTATAGGTA GACCAAATGT AGGGAAGTCA      60
ACATTTGTTA ATAGAGTGAT TGGCCATAAA ATAGCAATCA TGTCCGATAA AGCTCAAACA      120
ACTAGAAATA AAATTCAAGG TGTATGACA AGAGATGACG CGCAAATTAT ATTCAATGAT      180
ACGCCAGGTA TTCATAAACC TAAACACAAA TTAGGTGACT ATATGATGAA AGTCGCTAAA      240
AATACATTAT CTGAAATAGA TGCAATCATG TTTATGGTTA ATGCCAATGA GGAAATTGGA      300
CGAGGCGATG AATATATTAT AGAAATGTTG AAAAATGTTA AGACACCAGT ATTTTTAGTA      360
TTAAATAAAA TAGATTTAGT GCATCCAGAT GAATTAATGC CAAAGATTGA AGAATATCAA      420
AGTTATATGG ACTTTACAGA GATTGTACCT ATTTTCAGCAT TAGAAGGGCT AAATGTAGAT      480
CATTTTATTG ATGTTTTTAAA GACGTATTTA CCCGAAGGAC CTAAATATTA TCCAGATGAT      540
CAAATTTTCTG ACCATCCTGA ACAATTTGTA GTGGGTGAAA TCATTTCGTGA AAAAATCCTT      600
CATCTTACAA GTGAAGAAAT CCCTCATGCG ATTGGTGTTA ATGTGGACCG TATGGTTAAA      660
GAAAGCGAAG ATCGTGTTC TATCGAAGCA ACTATATATG TTGAAAGAGA TTCGCAAAAA      720
GGAATTGTCA TTGGAAGAGG CGGTAAAAAG TTAAGAGAAG TAGGAAAACG TGCGAGACGT      780
GATATAGAAA TGCTTCTAGG CTCTAAAGTA TACTTAGAAT TATGGGTCAA AGTTCAAAGA      840
GACTGGCGAA ACAAAGTTAA CTTTATTCGC CAAATTGGTT ATGTTGAAGA CCAAGATTAA      900

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 299 amino acids
- (B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Thr Glu His Lys Ser Gly Phe Val Ser Ile Ile Gly Arg Pro Asn
 1           5           10           15
Val Gly Lys Ser Thr Phe Val Asn Arg Val Ile Gly His Lys Ile Ala
      20           25           30
Ile Met Ser Asp Lys Ala Gln Thr Thr Arg Asn Lys Ile Gln Gly Val
      35           40           45
Met Thr Arg Asp Asp Ala Gln Ile Ile Phe Asn Asp Thr Pro Gly Ile
      50           55           60
His Lys Pro Lys His Lys Leu Gly Asp Tyr Met Met Lys Val Ala Lys
      65           70           75           80
Asn Thr Leu Ser Glu Ile Asp Ala Ile Met Phe Met Val Asn Ala Asn
      85           90           95
Glu Glu Ile Gly Arg Gly Asp Glu Tyr Ile Ile Glu Met Leu Lys Asn
      100          105          110
Val Lys Thr Pro Val Phe Leu Val Leu Asn Lys Ile Asp Leu Val His
      115          120          125
Pro Asp Glu Leu Met Pro Lys Ile Glu Glu Tyr Gln Ser Tyr Met Asp
      130          135          140
Phe Thr Glu Ile Val Pro Ile Ser Ala Leu Glu Gly Leu Asn Val Asp
      145          150          155          160
His Phe Ile Asp Val Leu Lys Thr Tyr Leu Pro Glu Gly Pro Lys Tyr
      165          170          175
Tyr Pro Asp Asp Gln Ile Ser Asp His Pro Glu Gln Phe Val Val Gly
      180          185          190
Glu Ile Ile Arg Glu Lys Ile Leu His Leu Thr Ser Glu Glu Ile Pro
      195          200          205
His Ala Ile Gly Val Asn Val Asp Arg Met Val Lys Glu Ser Glu Asp
      210          215          220
Arg Val His Ile Glu Ala Thr Ile Tyr Val Glu Arg Asp Ser Gln Lys
      225          230          235          240
Gly Ile Val Ile Gly Lys Gly Gly Lys Lys Leu Lys Glu Val Gly Lys
      245          250          255
Arg Ala Arg Arg Asp Ile Glu Met Leu Leu Gly Ser Lys Val Tyr Leu
      260          265          270
Glu Leu Trp Val Lys Val Gln Arg Asp Trp Arg Asn Lys Val Asn Phe
      275          280          285
Ile Arg Gln Ile Gly Tyr Val Glu Asp Gln Asp
      290          295

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 540 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

GGCACGAGCT AGTTTGGTAT CGGCAATTC TCAAGGATAC AGACCAGGTG ATTTTGAATC      60
AATAACTGTA ACCGTAGATG CAGATAAACC GTCATCACCT TGTGGTGCAT GTCGTCAAGT      120
TTTAAAGGAA TTATGTGATG ATGATATGCC TGTGTATATG ACAAATCATA AAGGAGATAT      180
GGTTATGATG ACAGNCGCAG AGTTACTACC ATTTGGATTT TCAGGAAAGG ATTTAGAATA      240
AATGACAGAA CATAAATCAG GATTTGTTTC AATTATAGGT AGACCAAATG TAGGGAAGTC      300
AACATTTGTT AATAGAGTGA TTGGCCATAA AATAGCAATC ATGTCCGATA AAGCTCAAAC      360
AACTAGAAAT AAAATTCAAG GTGTTATGAC AAGAGATGAC GCGCAAATTA TATTCAATGA      420
TACGCCAGGT ATTCATAAAC CTAAACACAA ATTAGGTGAT TATACTGATG AAAGTCGCTT      480
AAAATACATA TCTGAAATAG ATGCAATCAT GGTTTATGGC TAATTGCAAT GAGAAATTGG      540

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Thr Glu His Lys Ser Gly Phe Val Ser Ile Ile Gly Arg Pro Asn
 1             5             10             15
Val Gly Lys Ser Thr Phe Val Asn Arg Val Ile Gly His Lys Ile Ala
          20             25             30
Ile Met Ser Asp Lys Ala Gln Thr Thr Arg Asn Lys Ile Gln Gly Val
          35             40             45
Met Thr Arg Asp Asp Ala Gln Ile Ile Phe Asn Asp Thr Pro Gly Ile
          50             55             60
His Lys Pro Lys His Lys Leu Gly Asp Tyr Thr Asp Glu Ser Arg Leu
65             70             75             80

```

Lys Tyr Ile Ser Glu Ile Asp Ala Ile Met Val Tyr Gly
85 90

What is claimed is:

1. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
 - (a) a polynucleotide having at least a 70% identity to a polynucleotide encoding a polypeptide comprising amino acids 1 to 299 of SEQ ID NO:2;
 - (b) a polynucleotide which is complementary to the polynucleotide of (a); and
 - (c) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a) or (b).
2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.
3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.
4. The polynucleotide of Claim 2 comprising the nucleotide 1 to 900 set forth in SEQ ID NO:1.
5. The polynucleotide of Claim 2 comprising nucleotide 1 to 900 set forth in SEQ ID NO:1.
6. The polynucleotide of Claim 2 which encodes a polypeptide comprising amino acid 1 to 299 of SEQ ID NO:2.
7. An isolated polynucleotide comprising a member selected from the group consisting of:
 - (a) a polynucleotide having at least a 70% identity to a polynucleotide encoding the same mature polypeptide expressed by the ERA gene contained in NCIMB Deposit No. 40771;
 - (b) a polynucleotide complementary to the polynucleotide of (a); and
 - (c) a polynucleotide comprising at least 15 bases of the polynucleotide of (a) or (b).
8. A vector comprising the DNA of Claim 2.
9. A host cell comprising the vector of Claim 8.
10. A process for producing a polypeptide comprising: expressing from the host cell of Claim 9 a polypeptide encoded by said DNA.
11. A process for producing a cell which expresses a polypeptide comprising transforming or transfecting the cell with the vector of Claim 8 such that the cell expresses the polypeptide encoded by the cDNA contained in the vector.

12. A process for producing a ERA polypeptide or fragment comprising culturing a host of claim 9 under conditions sufficient for the production of said polypeptide or fragment.
13. A polypeptide comprising an amino acid sequence which is at least 70% identical to amino acid 1 to 299 of SEQ ID NO:2.
14. A polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:2.
15. An antibody against the polypeptide of claim 13.
16. An antagonist which inhibits the activity of the polypeptide of claim 13.
17. A method for the treatment of an individual having need of ERA comprising: administering to the individual a therapeutically effective amount of the polypeptide of claim 13.
18. The method of Claim 16 wherein said therapeutically effective amount of the polypeptide is administered by providing to the individual DNA encoding said polypeptide and expressing said polypeptide *in vivo*.
19. A method for the treatment of an individual having need to inhibit ERA polypeptide comprising: administering to the individual a therapeutically effective amount of the antagonist of Claim 16.
20. A process for diagnosing a disease related to expression of the polypeptide of claim 13 comprising:
determining a nucleic acid sequence encoding said polypeptide.
21. A diagnostic process comprising:
analyzing for the presence of the polypeptide of claim 13 in a sample derived from a host.
22. A method for identifying compounds which bind to and inhibit an activity of the polypeptide of claim 13 comprising:
contacting a cell expressing on the surface thereof a binding for the polypeptide, said binding being associated with a second component capable of providing a detectable signal in response to the binding of a compound to said binding, with a compound to be screened under conditions to permit binding to the binding; and
determining whether the compound binds to and activates or inhibits the binding by detecting the presence or absence of a signal generated from the interaction of the compound with the binding.

23. A method for inducing an immunological response in a mammal which comprises inoculating the mammal with ERA, or a fragment or variant thereof, adequate to produce antibody to protect said animal from disease.

24. A method of inducing immunological response in a mammal which comprises,
5 through gene therapy, delivering gene encoding ERA fragment or a variant thereof, for expressing ERA, or a fragment or a variant thereof *in vivo* in order to induce an immunological response to produce antibody to protect said animal from disease.

25. An immunological composition comprising a DNA which codes for and expresses a ERA polynucleotide or protein coded therefrom which, when introduced into a
10 mammal, induces an immunological response in the mammal to a given ERA polynucleotide or protein coded therefrom.

26. A polynucleotide consisting essentially of a DNA sequence obtainable by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in SEQ ID NO:1 under stringent hybridization conditions with a probe
15 having the sequence of said polynucleotide sequence set forth in SEQ ID NO:1 or a fragment thereof; and isolating said DNA sequence.

FIGURE 1. [SEQ ID NO:1]

```
1  ATGACAGAAC ATAAATCAGG ATTTGTTTCA ATTATAGGTA GACCAAATGT
51  AGGGAAGTCA ACATTGTGTA ATAGAGTGAT TGGCCATAAA ATAGCAATCA
101 TGTCCGATAA AGCTCAAAACA ACTAGAAATA AAATTCAAAG TGTATGACA
151 AGAGATGACG CGCAAATTAT ATTCAATGAT ACGCCAGGTA TTCATAAAACC
201 TAAACACAAA TTAGGTGACT ATATGATGAA AGTCGCTAAA AATACATTAT
251 CTGAAATAGA TGCAATCATG TTTATGGTTA ATGCCAATGA GGAAATTGGA
301 CGAGGCGATG AATATATTAT AGAAATGTTG AAAAAATGTTA AGACACCAGT
351 ATTTTITAGTA TTAAATAAAA TAGATTTAGT GCATCCAGAT GAATTAATGC
401 CAAAGATTGA AGAATATCAA AGTTATATGG ACTTTACAGA GATTGTACCT
```


Figure 1A

451 ATTCAGCAT TAGAAGGCT AAATGTAGAT CATTTATTG ATGTTTTTAAA
501 GACGTATTTA CCCGAAGGAC CTAAATATTA TCCAGATGAT CAAATTTTCAG
551 ACCATCCTGA ACAATTGTGTA GTGGGTGAAA TCATTCGTGA AAAAAATCCTT
601 CATCTTACAA GTGAAGAAAT CCTCATGCG ATTGGTGTTA ATGTGGACCG
651 TATGGTTAAA GAAAGCGAAG ATCGTGTTCA TATCGAAGCA ACTATATATG
701 TTGAAAGAGA TTCGCAGAAA GGAATTGTCA TTGGAAAAGG CGGTAAAAAAG
751 TTAAAAGAAG TAGGAAAACG TCCGAGACGT GATATAGAAA TGCTTCTAGG
801 CTCTAAAAGTA TACTTAGAAT TATGGGTCAA AGTCAAAAGA GACTGGCGAA
851 ACAAAGTTAA CTTTATTCGC CAAATTGGTT ATGTTGAAGA CCAAGATTAA

FIGURE 2. [SEQ ID NO:2]

1 MTEHKSGFVS IIGRPNVGKS TFVNRVIGHK IAIMSDKAQT TRNKIQGVMT
51 RDDAQIIFND TPGIHKPKHK LGDYMKVAK NTLSEIDAIM FMVNANEEIG
101 RGDEYIIEML KNVKTPVFLV LNKIDLVHPD ELMPKIEEYQ SYMDFTEIVP
151 ISALEGLNVD HFIDVLKTYL PEGPKYYPPDD QISDHPPEQFV VGEIIREKIL
201 HLTSEEIPHA IGVNVDRMVK ESEDRVHIEA TIYVERDSQK GIVIGKGGKK
251 LKEVGKRARR DIEMLLGSKV YLELWVKVQR DWRNKNVNFIR QIGYVEDQD

FIGURE 3. [SEQ ID NO:3]

GGCAGGAGCTAGTTTGGTATCGGCAATTCTCAAGGATACAGACCAGGTGATTTTGAATC
AATAACTGTAACCGTAGATGCAGATAAAACCGTCATCACCTTGTGGTGCATGTCGTCAAAGT
TTTAAAGGAATTATGTGATGATGATATGCCCTGTGTATATGACAAATCATAAAGGAGATAT
GGTTATGATGACAGNCGCAGAGTTACTACCAATTTGGATTTTCAGGAAAGGATTTAGAATA
AATGACAGAAACATAAAATCAGGATTTGTTTCAATTATAGGTAGACCACCAATGTAGGGAAGTC
AACATTTGTTAATAGAGTGATTGGCCATAAAAATAGCAATCATGTCCGATAAAGCTCAAAC
AACTAGAAATAAAATCAAGGTGTTATGACAAAGAGATGACGCGCAAAATTATATTCAATGA
TAGGCCAGGTATTCATAAAACCTAAACACAAATTAGGTGATTATCTGATGAAAAGTCGCCTT
AAAAATACATATCTGAAATAGATGCAATCATGGTTTATGGCTAATTGCAATGAGAAATTGG

FIGURE 4. [SEQ ID NO:4]

MTEHKSGFVSIIGRPNVGKSTFVNRVIGHKIAIMSDKAQTTRNKIQGVMTDRDDAQIIFND

TPGIHKPKHKLGDYTDESRLKYISEIDAIMVYG

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/02547

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12N 15/00, 15/31, 1/20; C07K 1/10, 14/195; A61K 39/00
US CL : 536/23.1; 435/252.3; 530/350; 424/184.1, 130.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. :

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

MEDLINE, BIOSIS, EMBASE, WPIDS, CAPLUS, APS, CURRENT BIOTECH ABS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	YAMASHITA, Y. et al. Molecular characterization of a Streptococcus mutants Mutants altered in environmental stress responses. J. Bacteriology. 1993, Vol.175, No. 19, pages 6220-6228, see entire document.	13-26
Y	PILLUTLA, C. R. et al. Cross-species complementation of the indispensable Escheria coli era gene highlits amino acid regions essential for activity. J. Bacteriology. 1995, Vol. 177, No. 8, pages 2194-2196, see entire document.	1-26
Y	Zuber, M. et al. Analysis of the rnc locus of Coxiella burnetti. Molecular Microbiology. 1994, Vol. 14, No. 2, pages 291-300, see entire document.	1-26

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

23 APRIL 1997

Date of mailing of the international search report

26 JUN 1997

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

KHALID MASOOD

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/02547

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LERNER, G. C. et al. Pleiotropic changes resulting from depletion of ERA, an essential GTP-binding protein in Escheria coli. Molecular Microbiology. 1991, Vol. 5, No. 4, pages 951-957, see entire document.	1-26
Y	SOOD, P. et al. Characterization of the autophosphorylation of Era, an essential Escheria coli GTPase. Molecular Microbiology. 1994, Vol. 12, No. 2, pages 201-208, see entire document.	1-26